SPURRING AI INNOVATION WITH SOUND DATA POLICY

The Impact of Data Policy on AI Development

At its core, Artificial Intelligence (AI) is a technology that augments human intelligence, helping people make better-informed decisions by identifying relationships, patterns, and trends in data that would be imperceptible to humans. Although AI research dates back several decades, advances in the availability of computing power, highly sophisticated algorithms, and data have recently accelerated its use in the marketplace.

AI systems are "trained" by ingesting enormous volumes of data. The benefits of AI are therefore dependent on the quantity and quality of data that is available for training. As a result, government policies affecting the ability to access and share data have a significant influence on the development of AI.

This issue snapshot discusses the relationship between sound government data policies and AI innovation.

1 Ensure Data Can Move Freely Across Borders

The free flow of data across borders is critical for the services that sustain global commerce, improve health and safety, promote social good, and enable the technologies of the future. In fact, data transfers are integral to every stage of the AI life cycle, from the development of predictive models to the deployment and use of AI systems.

The data used in AI systems often originates from many geographically dispersed sources, making it imperative that data can move freely across borders. Rules that limit cross-border data transfers invariably limit the insights and other benefits that AI systems can provide. For instance, the future of agriculture will be driven by data from drones and sensors that precisely measure the soil acidity, moisture retention, and fertility rates of every inch of a farmer's land. A cloud-based AI system can analyze this agricultural data to recommend real-time adjustments that improve crop yields while lowering the costs and environmental effects of farming. The ability for such a system to provide insights to farmers in remote regions of the world will, of course, depend on the free flow of data.

2 Access to Government Data and Public Sector Information

Government-generated data is a resource that can serve as a powerful engine for creating new jobs and promoting economic growth. At both the local and national level, governments collect and generate vast quantities of data that can be harnessed in the development of AI systems.

more >>

The future of agriculture will be driven by data from drones and sensors that precisely measure the soil acidity, moisture retention, and fertility rates of every inch of a farmer's land. A cloud-based AI system can analyze this agricultural data to recommend real-time adjustments that improve crop yields while lowering the costs and environmental effects of farming. The ability for such a system to provide insights to farmers in remote regions of the world will, of course, depend on the free flow of data.

» **Putting government-held data to use:** Sound data policies should ensure that any non-sensitive government-generated data asset is made freely available to the public in machine-readable formats. This data — a resource that would otherwise be unused — can improve services and lower prices. For instance, an AI system designed to improve supply chain efficiency might rely on government data about historical traffic flows, law enforcement event advisories, and weather patterns to recommend delivery routes that minimize congestion, reduce emissions, and improve public safety.

» **Preventing data lock-up through procurement:** Governments should carefully consider the effect that procurement policies can have on the availability of data. As a general matter, governments should avoid service agreements that would grant exclusive access or use rights to government datasets to any single private entity. Increasingly, government data is being generated by third-party vendors. For instance, local transit authorities may contract with third-party vendors to analyze data generated by buses and trains, or by sensors embedded in street lights and roadways. Governments contracting for such services should ensure that any statistical data created or maintained on its behalf as part of the agreement is not subject to access or use restrictions. Rather, data provided to governments as part of such procurement contracts should be treated like any other government data asset and should be made freely accessible for public use.

3 Facilitate Value-Added Data Services

Data is the key to growth in all sectors of the economy, but each piece of data has little inherent value. It is only when used as an input to other value-added services, such as AI, that it contributes to the projected $15 trillion addition to global GDP by 2030.[1]

Policies that artificially increase the costs for acquiring the data used to train AI systems will ultimately increase the costs of these technologies for customers and decrease the incentive to develop and use new technology — potentially reducing overall consumer welfare.

Government-generated data is a resource that can serve as a powerful engine for creating new jobs and promoting economic growth. At both the local and national level, governments collect and generate vast quantities of data that can be harnessed in the development of AI systems.

Governments should pursue policies that facilitate the business-to-business exchange of data and boost the development of AI services, including by:

» Ensuring companies can enter enforceable contracts that create data sharing arrangements;

» Avoiding the creation of new rights in business data that could add unnecessary transaction costs; and

» Allowing companies to freely perform data analytics, including text and data mining, on any content to which they have lawful access.

4 Maintain Predictable Competition Policies

Fierce competition among AI providers is creating dynamic efficiencies that push the creation of new services. This benefits every sector of the economy — helping companies optimize their manufacturing processes, improve their supply chains, secure their networks, and enhance their products and services. These downstream benefits have been driven by a diverse set of AI companies, with small- and medium-sized firms playing an important role.[2]

Policymakers should avoid creating AI-specific competition rules, such as compulsory licensing of data sets. Creating and structuring data sets is resource intensive. A regulatory prohibition on maintaining the exclusive benefits of that investment will deter the investment and impede the development of AI.

AI remains a burgeoning field with quickly evolving market dynamics. The existing framework of competition law is intentionally designed to be flexible enough to address technological innovations, and is better suited than ex-ante regulatory intervention.

[1] BSA, *What's the Deal with Big Data,* available at http://data.bsa.org/wp-content/uploads/2015/12/bsadatastudy_en.pdf.

[2] There are currently more than 2,000 AI startups that have raised nearly $30 billion in funding. See Vala Afshar, "AI Is Transformational Technology and Major Sector Disruptor," *Huffington Post* (December 5, 2017), available at https://www.huffingtonpost.com/entry/ai-is-transformational-technology-and-major-sector_us_5a259dbfe-4b05072e8b56b6e.